Published in final edited form as:
Nat Genet. 2015 October ; 47(10): 1121–1130. doi:10.1038/ng.3396.

A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease

A full list of authors and affiliations appears at the end of the article.
[#] These authors contributed equally to this work.

Abstract

Existing knowledge of genetic variants affecting risk of coronary artery disease (CAD) is largely based on genome-wide association studies (GWAS) analysis of common SNPs. Leveraging phased haplotypes from the 1000 Genomes Project, we report a GWAS meta-analysis of 185 thousand CAD cases and controls, interrogating 6.7 million common (MAF>0.05) as well as 2.7 million low frequency (0.005<MAF<0.05) variants. In addition to confirmation of most known CAD loci, we identified 10 novel loci, eight additive and two recessive, that contain candidate genes that newly implicate biological processes in vessel walls. We observed intra-locus allelic heterogeneity but little evidence of low frequency variants with larger effects and no evidence of synthetic association. Our analysis provides a comprehensive survey of the fine genetic architecture of CAD showing that genetic susceptibility to this common disease is largely determined by common SNPs of small effect size.

Coronary artery disease (CAD) is the main cause of death and disability worldwide and represents an archetypal common complex disease with both genetic and environmental determinants[1,2]. To date, 48 genomic loci have been found to harbour common SNPs in genome-wide significant association with the disease. Previous GWAS of CAD have tested the common disease/common variant hypothesis with meta-analyses typically based on HapMap imputation training sets or tagging SNP arrays with up to 2.5 million SNPs (85% with MAF > 0.05)[3,4]. The 1000 Genomes Project[5] has considerably expanded the coverage of human genetic variation especially for lower frequency and insertion/deletion variants (indels). We assembled 60,801 cases and 123,504 controls from 48 studies for a GWAS meta-analysis of CAD; 34,997 (57.5%) of the cases and 49,512 (40.1%) of the controls had been previously included in our Metabochip-based CAD meta-analysis (Supplementary Fig. 1)[3]. Imputation was based on the 1000 Genomes phase 1 version 3 training set with 38 million variants of which over half are low frequency (MAF < 0.005) and one-fifth are

Correspondence should be addressed to H.W (hugh.watkins@rdm.ox.ac.uk), S.Kathiresan (sekar@broadinstitute.org), R.M. (rmcpherson@ottawaheart.ca), M.F. (martin.farrall@well.ox.ac.uk).
[133]These authors jointly supervised this work.

URLs The Ensembl database, http://www.ensembl.org; the University of Chicago eQTL browser, http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl; the GTEx Portal, http://www.gtexportal.org/home; the Geuvadis Data Browser, http://www.ebi.ac.uk/Tools/geuvadis-das; the CARDIoGRAMplusC4D Consortium website, http:// http://www.cardiogramplusc4d.org

Competing financial interests
The authors declare no competing financial interests.

common (MAF > 0.05) variants. The majority (77%) of the participants were of European ancestry; 13% and 6% were of south (India and Pakistan) and east (China and Korea) Asian ancestry with smaller samples of Hispanic and African Americans (Supplementary Table 1). Case status was defined by an inclusive CAD diagnosis (e.g. myocardial infarction (MI), acute coronary syndrome, chronic stable angina, or coronary stenosis >50%). After selecting variants that surpassed allele frequency (MAF > 0.005) and imputation quality control criteria in at least 29 (>60%) of the studies, 8.6 million SNPs and 836K (9%) indels were included in the meta-analysis (Fig. 1); of these, 2.7 million (29%) were low frequency variants (0.005 < MAF < 0.05).

Scanning for additive associations

The results of an additive genetic model meta-analysis are summarized in Manhattan plots (Fig. 2 and Supplementary Fig. 2). 2,213 variants (7.6% indels) showed significant associations ($P < 5 \times 10^{-8}$) with CAD with a low false discovery rate (FDR q-value < 2.1×10^{-4}). When these 2,213 variants are grouped into loci, eight represent regions not previously reported at genome-wide levels of significance (Fig. 2; Table 1). Of 48 previously reported loci at genome-wide levels of significance, 47 showed nominally significant associations (Supplementary Table 2). The exception was rs6903956, the lead SNP for the *ADTRP–C6orf105* locus detected in Han Chinese[6], which previously showed no association in the Metabochip meta-analysis of Europeans and South Asians[3]. Thirty-six previously reported loci showed genome-wide significance (Supplementary Table 2). Monte Carlo simulations, guided by published effect-sizes, suggest that our study was powered to detect 34 of the previously reported loci (95%CI 31 – 41 loci) at genome-wide significance. Hence, our findings are fully consistent with the previously identified CAD loci. The majority of the loci showing GWAS significance in the present analysis were well imputed (82% with imputation quality > 0.9) (Fig. 3a) and had small effect sizes (odds ratio (OR) < 1.25) (Fig. 3b). An exception was the lead SNP in the novel chromosome 7q36.1 (*NOS3*) locus, rs3918226, which was only moderately well imputed (quality 0.78) but the validity of this association is supported by existing genotype data as it was present on the HumanCVD BeadChip available for some of the cohorts used in the present analysis and therefore directly measured genotypes could be compared with imputed genotypes (Supplementary Table 3)[7]. Three additional lower frequency and moderately well imputed SNPs in *LPA* and *APOE* (Fig. 3a), which were not previously reported in CAD GWAS[3,4], also showed strong associations (*LPA*: rs10455872 $P = 5.7 \times 10^{-39}$, rs3798220 $P = 4.7 \times 10^{-9}$; *APOE*: rs7412 $P = 8.2 \times 10^{-11}$). The *LPA* SNPs have been previously shown to be strongly associated with CAD in candidate gene studies based on experimental genotype data[7,8]. SNP rs7412 encodes the epsilon 2 allele of *APOE* and it has been well documented that ε2 carriers have lower cholesterol levels and significant protection from CAD was confirmed in a large meta-analysis[9] and in the Metabochip study ($P = 0.0009$)[3]. However, rs7412 is not present on most commercially available genome-wide genotyping arrays and cannot be imputed using HapMap reference panels, supporting the value of the expanded coverage of the 1000 Genomes reference panels. Finally, SNP rs11591147, which encodes the low frequency (MAF = 0.01) R46L variant in *PCSK9* that has been associated with low LDL cholesterol levels and cardioprotection[10-13], was imperfectly imputed (quality = 0.61). Nonetheless

these data provide the strongest evidence to date for a protective effect of this variant for CAD ($P = 7.5 \times 10^{-6}$).

Scanning for non-additive associations

Few GWAS of CAD have systematically scanned for associations that include dominance effects and few truly recessive loci have been reported[14,15]. We used a recessive inheritance model to search for susceptibility effects conferred by homozygotes for the minor (i.e. less frequent) allele. Two novel recessive loci were identified with MAF 0.09 and 0.36 and genotypic OR 0.67 and 1.12 (Table 1, Fig. 2); these loci showed very little evidence of association under an additive model (Table 1). A supplementary dominant model analysis revealed multiple strong associations with variants that all overlapped with loci identified in the additive model analysis (Supplementary Table 4).

MI sub-phenotype analysis

Sub-group analysis in cases with a reported history of MI (~ 70% of total number of cases) did not identify any additional associations reaching genome-wide significance. The association results for the MI sub-phenotype for the 48 previously known CAD loci and the 8 novel additive CAD loci discovered in this study are shown in Supplementary Table 5. Supplementary Fig. 3 compares the OR for the lead SNPs at 56 loci for the broader CAD phenotype (full cohort) and the MI sub-phenotype. While, as expected, for most of the loci the ORs are very similar, for *ABO* and *HDAC9* the ORs are sufficiently distinct for their 95% confidence intervals to lie away from the line of equality, suggesting that *ABO* preferentially associates with MI and *HDAC9* with stable coronary disease but not infarction *per se*.

FDR and heritability analysis

We performed a joint association analysis to search for evidence of synthetic associations[16] where multiple low-frequency susceptibility variants at a locus might be in LD with a common variant that is discovered as the lead variant in a GWAS, and to compile an FDR-defined list of informative variants for annotation and heritability analysis[3]. Variants that showed suggestive additive associations ($P < 5 \times 10^{-5}$) were assigned to 214 putative susceptibility loci centred on a lead variant ±1 centiMorgan and all variants in these loci examined; consequently the search space for the joint analysis included 1,399,533 variants. Using the GCTA software[17] to perform an approximate joint association analysis (**Online Methods**), we identified 202 variants (q-value < 0.05) in 129 loci (Supplementary Table 6) with multiple (2 – 14) tightly linked variants in 57% of CAD loci. The 202 FDR variants were mostly common (median MAF 0.22) and well imputed (median imputation quality 0.97). 95 variants (explaining 13.3±0.4% of CAD heritability) mapped to 44 GWAS significant loci. 93 variants (explaining 12.9±0.4% of CAD heritability) mapped to loci that include a previously reported GWAS significant variant; 109 variants (explaining a further 9.3±0.3% of CAD heritability) mapped to other loci. Fifteen low frequency (MAF < 0.05) variants explained only 2.1±0.2% of CAD heritability noting that our study was ~90% powered to detect OR > 1.5 with low frequency variants (Supplementary Table 7).

Common variants showing typical GWAS signals might be coupled with one or more low frequency variants with relatively large effects[16]. We found no evidence for such synthetic associations in the joint association analysis i.e. all low frequency variants were either lead variants or jointly associated (q-value < 0.05) with a common variant. Twenty of the 202 FDR variants (9.9%) were indels (4–14 bp size) compared to 8.8% of all the variants in the meta-analysis ($P = 0.60$). Low frequency variants (MAF < 0.05) were strikingly underrepresented (6.9% vs. 29.0%; $P = 4.9 \times 10^{-12}$) which may reflect on the statistical power to detect modest effects associated with these variants (*vide infra*).

Annotation and ENCODE analysis

Functional annotations were assigned to the 9.4 million variants studied in the CAD additive meta-analysis using the ANNOVAR software[18] (Supplementary Table 8). The 202 FDR variants were depleted in intergenic regions ($P = 2.5 \times 10^{-7}$) and enriched in introns ($P = 0.00035$). Variants were also assigned to three sets of ENCODE (Encyclopedia of DNA Elements) features, namely histone/chromatin modifications (HM), DNase I hypersensitive sites (DHS) and transcription factor binding sites (TFBS) (Supplementary Table 9). The FDR variants showed independent enrichment across 11 cell types for the HM ($P = 2.8 \times 10^{-6}$) and DHS ($P = 0.0003$) ENCODE feature sets and with genic annotation status ($P = 0.0013$) (Supplementary Table 10 and Supplementary Table 11). These associations were also evident in three cell types selected with maximal CAD relevance with a 2.6-fold enrichment for DHS, 2.2-fold enrichment for HM and 1.6-fold enrichment for genic status (Supplementary Table 12 and Supplementary Table 13). These findings suggest that the 202 FDR variants are enriched for functional variants with potential relevance to CAD pathogenesis.

Post hoc power calculations

8.2M of 9.4M (87%) of the analysed variants were highly powered (> 90%) to detect an OR ≥ 1.3 (Supplementary Table 7). The number of variants with power ≥ 90% to detect associations varies systematically with allele frequency and imputation quality (Supplementary Fig. 4 for OR = 1.3); 1.5M of 2.7M (55%) low frequency variants (0.005 < MAF < 0.05) in the meta-analysis were adequately powered to detect an OR ≥ 1.3 as most of these variants were accurately imputed (median imputation quality 0.94, interquartile range 0.88, 0.98). With more common variants (MAF > 0.05) almost all (99.8%) were highly powered to detect OR ≥ 1.3. However, in terms of the total coverage of low frequency variation, only 15.3% of 9.3M low frequency (0.005 < MAF < 0.05) variants in the 1000 Genomes phase 1 v3 training set surpassed the allele frequency and imputation quality entry criteria in the 60% of the studies required for inclusion in the meta-analysis and were predicted to be adequately powered to detect significant associations; 100% of these variants were highly powered (>90%) to detect OR ≥ 3.15.

Interrogation of 10 novel additive and recessive loci

We examined whether there were any eQTLs, associations with known cardiovascular risk factors or prior evidence of involvement of genes with atherosclerotic processes in each of

the newly identified loci to define putative mechanisms by which the loci may affect risk of CAD.

Chromosome 4q12 (*REST – NOA1*) locus: The lead SNP rs17087335 lies within an intron of the nitric oxide associated 1 (*NOA1*) gene; 23 SNPs in linkage disequilibrium (LD r^2>0.8) show CAD associations ($P < 1 \times 10^{-6}$) across the *NOA1* and the repressor element-1 silencing transcription factor (*REST*) genes (Fig. 4a). *NOA1* is a GTP binding protein involved in the regulation of mitochondrial respiration and apoptosis[19]. *REST* is a transcription factor that suppresses the expression of voltage dependent sodium and potassium channels[20] and has been shown to maintain vascular smooth muscle cells in a quiescent non-proliferative state and is itself down-regulated in neointimal hyperplasia[21]. SNP rs17087335 shows a cis-eQTL signal with *REST* in lung[22] (Supplementary Table 14).

Chromosome 7q36.1 (*NOS3*) locus: The lead SNP rs3918226 (MAF = 0.07) lies in the first intron of the nitric oxide synthase 3 (*NOS3*) gene (Fig. 4b). This SNP has been tentatively associated with CAD (OR = 1.14, $P = 1.4 \times 10^{-4}$) in a candidate gene meta-analysis based on 15.6K cases and 35K controls genotyped with the HumanCVD BeadChip[7] and firmly associated with essential hypertension (OR = 1.34, $P = 1.0 \times 10^{-14}$)[23]. *NOS3* is involved in production of nitric oxide (NO), a potent vascular smooth muscle relaxant and a well-studied candidate gene for CAD. Indeed, proteins forming the NO receptor (soluble guanylyl cyclase) display both linkage as well as genome-wide association with CAD[3,24]. There are several overlapping ENCODE features in the *NOS3* intron 1 suggesting a functional role for rs3918226. However, there are 30 genes neighbouring *NOS3* within a ±1 centiMorgan window and the current data does not allow candidacy of these genes to be excluded. A non-synonymous SNP, rs1799983, in *NOS3* previously associated with cardiovascular phenotypes[25] is in weak LD with rs3918226 but did not achieve significance in an additive or joint association analysis.

Chromosome 11p15.4 (*SWAP70*) locus: SNP rs10840293 is intronic to the switch-associated protein-70 gene (*SWAP70*) (Fig. 4c). SWAP-70 is a signalling molecule involved in the regulation of filamentous-actin networks[26] in cell migration and adhesion. SNP rs10840293 with other SNPs in strong LD are cis-eQTLs in naïve and challenged monocytes[27], SNP rs93138 shows a strong association with CAD ($P = 5.5 \times 10^{-8}$) and is a cis-eQTL in naïve and challenged monocytes[28], fat[29], skin[29] and lung[22] (Supplementary Table 14); three of these SNPs (rs93138, rs173396 and rs472109) are intronic and lie within ENCODE regulatory functional elements. Although this CAD-associated locus includes 33 genes, the eQTL and ENCODE data implicate *SWAP70* as a plausible causal gene and suggest putative causal SNPs.

Chromosome 15q22.33 (*SMAD3*) locus: The lead SNP rs56062135 is intronic to *SMAD3* and the CAD association is tightly localized between two recombination hot spots (Fig. 4d). Mice lacking Smad3, a major downstream mediator of TGF-β, show enhanced neointimal hyperplasia with decreased matrix deposition in response to vascular injury[30]. *SMAD3* had been tentatively associated in an earlier CAD GWAS[31], although that lead SNP (rs17228212) is in linkage equilibrium with rs56062135, showed modest association (P = 0.009) in the present GWAS and no evidence of joint association (Supplementary Table 6).

Chromosome 15q26.1 (*MFGE8 – ABHD2*) locus: The lead intergenic SNP rs8042271 maps 117kb upstream of milk fat globule-EGF factor 8 (*MFGE8*) and 57kb upstream from abhydrolase domain-containing protein 2 (*ABHD2*) (Fig. 4e). *MFGE8* (lactadherin) has a crucial role in VEGF dependent neovascularization[32] and is secreted from activated macrophages and binds to apoptotic cells, facilitating phagocytic engulfment[33]. *ABHD2*[34] has been shown to be expressed in human atherosclerotic lesions with higher levels in patients with unstable angina. There were no overlapping risk factor QTL, eQTL or ENCODE features to guide the nomination of a putative causal gene.

Chromosome 17q23.2 (*BCAS3*) locus: The lead intronic SNP rs7212798 lies in the breast carcinoma amplified sequence 3 (*BCAS3*) gene (Fig. 4f). Multiple variants in LD with rs7212798 map to *BCAS3* introns and show strong association with CAD. *BCAS3* encodes the Rudhira protein that has been shown to activate Cdc42 to affect actin organization and control cell polarity and motility in endothelial cells thus contributing to angiogenesis[35].

Chromosome 18q21.32 (*PMAIP1 – MC4R*) locus: The lead intergenic SNP rs663129 lies 266 kb downstream of the phorbol-12-myristate-13-acetate-induced protein 1 (*PMAIP1*) gene and 200kb downstream of the melanocortin 4 receptor (*MC4R*) gene (Fig. 4g). *PMAIP1* is a HIF1A-induced proapoptotic gene that mediates hypoxic cell death by the generation of reactive oxygen species[36]. *MC4R* is a well-studied obesity-related locus and the variant (and corresponding proxy variants) that were associated with higher CAD risk were also associated with BMI ($P = 6\times10^{-42}$) and obesity-associated risk factors including higher triglyceride and lower HDL concentrations and type 2 diabetes[37-41]. However, we found no eQTL data or ENCODE features for the lead or proxy SNPs to further implicate *MC4R* as the causal gene underlying the CAD susceptibility.

Chromosome 22q11.23 (*POM121L9P – ADORA2A*) locus: The lead SNP rs180803 lies in the non-coding RNA POM121 transmembrane nucleoporin-like 9, pseudogene (*POM121L9P*). A ±1 centiMorgan region spans 1.2Mb and includes 21 variants that are genome-wide significantly associated with CAD, most of them in LD ($r^2>0.6$) with the lead SNP and mapping to intronic regions of the *SPECC1L* and *ADORA2A* genes (Fig. 4h).

Chromosome 12q24.23 (*KSR2*) locus: The lead SNP rs11830157 (MAF=0.36) associates with CAD risk in a recessive model (genotypic OR = 1.12) and is intronic to the kinase suppressor of ras 2 (*KSR2*) gene (Fig. 4i) and overlaps with ENCODE functional elements. *KSR2* interacts with multiple proteins including AMPK and rare loss-of-function coding variants in *KSR2* are associated with severe obesity, hyperphagia and insulin resistance, a phenotype recapitulated in Ksr2 null mice[42].

Chromosome 19q13.11 (*ZNF507 – LOC400684*) locus: The lead SNP rs12976411 (MAF=0.09) lies in an uncharacterized non-coding RNA (*LOC400684*) and is 3.4 kb 3’ of *ZNF507* (Fig. 4j). The minor allele shows a protective CAD effect (genotypic OR = 0.69) in the recessive model. ENCODE analysis of this locus suggests that several SNPs including rs12981453 and rs71351160, which are in strong LD ($r^2>0.8$) and intronic to *ZNF507* overlap with ENCODE functional elements.

Discussion

We demonstrate that the ability of GWAS to investigate the genetic architecture of complex traits is enhanced by the 1000 Genomes Project. This has allowed us to conclude that low frequency variants of larger effect, synthetic associations and insertion/deletion polymorphisms are unlikely to explain a significant portion of the missing heritability of CAD. Rather, all ten novel CAD associations identified in the present analysis, as well as all but one of the previously identified loci, are represented by risk alleles with a frequency of greater than 5%. Thus, this comprehensive analysis strongly supports the common disease/common variant hypothesis[43] given that it was powered to detect MAF < 0.05 variants with OR > 1.5. Moreover, risk alleles are significantly clustered within or close to genes and enriched in regions with functional annotations. Finally, genes implicated by this unbiased approach suggest hypotheses that explore the biology of the arterial vessel wall as a critical component of CAD pathogenesis.

The success of the GWAS meta-analysis strategy to map common, small-effect susceptibility variants for complex diseases has leant heavily on genotype imputation with publically available training sets. The 1000 Genomes Project provides a substantial step-up from the HapMap era in terms of coverage of lower frequency variants and the integration of insertion/deletion polymorphism (Fig. 1). Lead SNPs for 4 of the 10 novel CAD loci were either absent or imperfectly tagged ($r^2 < 0.8$) in the HapMap2 training set, which reduced the power of discovering these loci in previous GWAS meta-analyses. Although lower frequency variants often show geographical differentiation[5], the phase1 v3 training set includes numerous low MAF variants that are tractable to a global meta-analysis that includes multiple continental ancestry groups. Key SNPs in apolipoproteins E and (a) and *PCSK9*, which mediate their CAD-effects via LDL-cholesterol linked mechanisms showed strong associations and reinforces the sensitivity of our 1000 Genomes analysis to detect lower frequency, imperfectly imputed susceptibility variants which were missed in HapMap-based GWAS.

Association analysis under the customary additive inheritance model widely used in GWAS analyses is optimally powered to detect traits with no dominance variance but conveniently has adequate power to detect dominantly inherited traits as well[44]. However, the additive model is systematically underpowered to detect recessively inherited traits particularly for lower frequency alleles[44]. This motivated our recessive meta-analysis, which revealed two novel CAD loci, *KSR2* and *ZNF507-LOC400684* that escaped detection in a conventional additive association scan.

Our GWAS explores two potential sources of missing heritability for CAD as it includes indels and an extended panel of lower frequency variants. Although there was no evidence that indels were systematically enriched for CAD associations, they were present in 10% of the 202 variants with an FDR q-value < 5%. In terms of surveying the totality of human genetic variation, the 1.5M of 2.7M lower frequency variants that were included in the meta-analysis with power to detect moderate-penetrance alleles (OR > 1.3) might seem modest. Yet the relative paucity of significant associations, 15 variants with MAF < 0.05 that explain 2% of CAD heritability and provide no evidence of synthetic associations, will temper

expectations for the role of low frequency variants underlying CAD susceptibility, specifically with respect to risk prediction in a population-based setting. It is though important to acknowledge that GWAS based on SNP array data has limited power to resolve genes carrying rare mutation burdens. For example, *LDLR*[45], *APOA5*[45], *APOC3*[46] and *NPC1L1*[47] are loaded with susceptibility or protective mutations for CAD. These discoveries were only revealed by whole-exome sequencing studies in large series of cases and controls yet explain less than 1% of the missing heritability of CAD[45].

Annotation analysis showed that the CAD associated variants were significantly clustered within or close to genes. Furthermore, there was a strong and independent enrichment for an overlap of the CAD associations with ENCODE features, particularly in cell types relevant to CAD pathogenesis. This phenomenon has been previously reported for other diseases and traits[48] and can guide candidate gene nomination and the design of future functional studies. We found few suggestions of overlap with risk factor QTL or eQTL in available datasets; this may in part reflect that the use of proxy variants can be limiting in cross-referencing 1000 Genomes and HapMap association databases.

Coronary atherosclerosis underlies the development of the vast majority of cases of MI; therefore the two are intimately related. However, additional factors such as plaque vulnerability or the extent of the thrombotic reaction to plaque disruption may predispose to MI in the presence of CAD[49]. We confirmed that *ABO* is particularly associated with risk of MI[50] suggesting that this locus may specifically increase the risk of plaque rupture and/or thrombosis. In contrast, *HDAC9* showed a stronger association with CAD than with MI suggesting that it might predispose to atherosclerosis but not the precipitant events leading to an MI. However, *HDAC9* shows even stronger associations with ischemic strokes involving thrombosis or embolism due to atherosclerosis of a large artery[51]. Although further epidemiological as well as experimental data are required to substantiate these findings, they suggest that certain loci may affect distinct mechanisms related to the development and progression of CAD.

Several of the genes implicated to date in large-scale analyses of CAD susceptibility encode proteins with a known role in the biology of circulating risk factors for CAD, notably lipid levels and the metabolism of lipoproteins; others relate to other known atherosclerosis risk factors, such as genes implicated in systemic inflammation and in hypertension. This is unsurprising, partly because of the undoubted importance of these known risk factors in the etiology of CAD but also because some of these prior analyses particularly targeted genes involved in risk factor traits, for example the HumanCVD BeadChip[52] was based on candidate genes and the Metabochip[3,53] drew on prior association data with risk factor traits as well as an earlier HapMap2 CAD GWAS meta-analysis[54]. The current experiment adopts a completely unbiased approach and is the first to do so at very large scale. In this respect, it is notable that for some of the novel loci where genomic data, biological precedent and eQTL associations suggest a plausible novel CAD gene, the genes so implicated have well documented roles in vessel wall biology. Their gene products are involved in diverse processes including cell adhesion and leukocyte and vascular smooth muscle cell (VSMC) migration (*SWAP70*[26] and *ABHD2*[55]), VSMC phenotypic switching (*REST*[20]), TGF-β

signaling (*SMAD3*[56,57]), anti-inflammatory and infarct sparing effects (*ADORA2A*[58] and *MFGE8*[59]), angiogenesis (*BCAS3*[35]) and nitrous oxide signaling (*NOS3*[24]).

It is important to note that these putative new susceptibility genes require substantial further investigation and validation before firm vascular biology links can be established. A number of preventative strategies target the vessel wall (control of blood pressure, smoking cessation) but the large majority of existing drug treatments for lowering CAD risk operate through manipulating circulating lipids and few directly target vessel wall processes. Detailed investigation of new aspects of vessel wall biology that are implicated by genetic association, but have not previously been explored in atherosclerosis, may provide new insights into its complex aetiology and hence new targets.

Online Methods

Association Analysis

Three models of heritable disease susceptibility were analysed by logistic regression 1) an additive model where the log(genotype risk ratio) (log(GRR)) for a genotype was proportional to the number of risk alleles, 2) a recessive model where the log(GRR) for homozygotes for the minor allele was compared with a reference risk in pooled heterozygotes and homozygotes for the major allele and 3) a dominant model where the log(GRR) for homozygotes for the minor allele pooled with heterozygotes was compared with a reference risk in homozygotes for the major allele. Minor and major alleles were identified by reference to the allele frequencies in the pooled populations (i.e. all continents) of 1000 Genomes phase 1 v3 data. For the recessive and dominant analyses, genotype probabilities to allow for variable imputation quality were analysed by all contributing studies; for the additive analysis, genotype probabilities or allelic dosages were used (Supplementary Table 1).

Data Quality Control

Association data for each contributing study were individually filtered by MAF > 0.005 (estimated in combined cases and controls) and an imputation quality metric, rsq > 0.3 for minimac or info_proper > 0.4 for IMPUTE2[61]. Allele frequencies for each study were binned and compared with other studies to detect systematic flipping of alleles (Supplementary Fig. 5). Over-dispersion of association statistics was assessed by the genomic control method[62] (Supplementary Table 15) and adjusted values were submitted for meta-analysis. Variants that were retained in at least 60% of the studies were submitted for meta-analysis using the GWAMA program[63]. Following an inverse-variance weighted fixed-effects meta-analysis, heterogeneity was assessed by Cochran's Q statistic[64] and the I^2 inconsistency index[65] and variants showing marked heterogeneity were reanalysed using a random-effects model[66]. Over-dispersion in the resulting meta-analysis statistics was adjusted for by a second application of the genomic control procedure (Supplementary Fig. 6).

FDR estimation

The false discovery rate (FDR) was assessed using a step-up procedure encoded in the *qqvalue* Stata program[67]. This procedure has been reported to be well controlled under positive regression-dependency conditions[68]; simulations based on 1000 permuted replicates of the PROCARDIS imputed data demonstrated that the FDR was conservatively controlled (theoretical q-value = 0.05, empirical q-value = 0.026 95%CI 0.017 – 0.038) in the context of the linkage disequilibrium patterns prevalent in the 1000 Genomes phase 1 v3 training set.

GCTA and heritability analysis

Joint association analysis of the CAD additive meta-analysis results was performed using the GCTA software[17] that fits an approximate multiple regression model based on summary association statistics and linkage disequilibrium information derived from a reference genotype database (here the 1000 Genomes phase 1 v3 training set for all continents/ populations that includes genotypes for 1,092 individuals). In this analysis, the lead variant is not necessarily retained in the final joint association model in situations where there might be multiple associated variants in strong LD. The accuracy of this analysis depends on appropriate ancestry matching as well as the sample size of the reference genotype panel to ensure that estimated LD correlations are unbiased and acceptably precise[69]. Simulations suggest that the expected correlation between p-values based on the GCTA method using a reference panel of 1,000 genotyped samples and p-values from an “exact” multiple regression based on experimental genotypes will be between 0.90 and 0.95[69]. We investigated the empirical accuracy of the GCTA joint association analysis by comparing GCTA joint association results with a standard multiple logistic regression analysis in 4 contributing studies (Supplementary Fig. 7). This showed that 95% of the betas and SE were accurately approximated. The $-\log_{10}$ p-values from the two analyses were positively correlated ($0.86 < \text{rho} < 0.93$) with the GCTA method showing an insignificant trend ($P > 0.20$) to yield slightly inflated values.

Heritability calculations were based on a multifactorial liability threshold model[70] assuming that the disease prevalence was 5% and that the total heritability of CAD was 40%[3]; multiple regression estimates of allelic effect sizes were used following the GCTA joint association analysis. Standard errors of the heritability estimates were estimated by Monte Carlo sampling with 1,000 replicates (i.e. for each variant, draw effect sizes (beta) randomly from the variants’ beta±SE estimate, calculate heritability for each beta-by-replicate draw, sum heritability across N variants within each replicate and finally calculate the standard error of the heritability estimates across the 1,000 replicates).

Power Calculations

The power to detect genetic associations depends on the magnitude of the genetic risk (i.e. effect size), the type 1 error rate, the risk allele frequency and imputation quality and the sample size. Non-centrality parameter calculations were based on double genomic controlled standard error estimates from the additive model meta-analysis; these estimates integrate information on allele frequency, imputation quality and sample size, which typically vary across studies. The type 1 error was set at 5×10^{-8} and an additive risk model was assumed.

Risk Factor QTL survey

The 10 novel CAD associated loci were scanned for associations with CAD heritable risk factors using publically available resources including large-scale GWAS consortia data downloads[37-41,71-73] and the NHGRI GWAS catalogue accessed May 2014[74]. As previous risk factor GWAS were mainly based on HapMap2 imputed datasets, all SNPs in LD (r^2>0.8 based on 1000 Genomes phase 1 v3 ALL reference panel) with the novel variants were examined for risk factor associations. The novel associated loci were cross-referenced with known *cis* and *trans* eQTL associations from the University of Chicago eQTL browser (accessed July 2014), the GTEx Portal (accessed June 2014), the Geuvadis Data Browser (accessed June 2014) and other published data[22,28,29,75-79].

Annotation and ENCODE analysis

Variants were annotated using the ANNOVAR[18] (version Aug 2013) software based on a GRCh/hg19 gene annotation database. Upstream/downstream status was assigned to variants that mapped ≤1kb from the transcript start/end. Variants without intergenic annotation were assigned a genic annotation status (42%). Supplementary Table 8 shows the annotation status of 9.4M variants included in the CAD additive meta-analysis; 86% of the genic variants map to introns.

ENCODE features were downloaded from the Ensembl database using the Funcgen Perl API module release 75. The list of the ENCODE experiments stored in the Ensembl database can be browsed at http://Feb2014.archive.ensembl.org/Homo_sapiens/Experiment/Sources?db=core;ex=project-ENCODE-. This summarized 100 different types of functional evidence in 11 different cell types, a total of 379 ENCODE experiments that revealed 6,099,034 features. Variants that overlaid one or more of these features were cross-tabulated with their ANNOVAR annotation status (Supplementary Table 10); 50% of variants mapped to one or more ENCODE features and variants in ENCODE features were strongly enriched for genic annotation status. Variants were grouped into three functional sets, histone/chromatin modifications (HM), DNase I hypersensitive sites (DHS) and transcription factor binding sites (TFBS) (Supplementary Table 9). Cell types were grouped into CAD relevant and others (Supplementary Table 12) based on their potential roles in CAD pathophysiology; hepatocytes (e.g. lipid metabolism[80]), vascular endothelial cells (atherosclerosis[81]) and myoblasts (injury/repair[82]) were selected as being most relevant to the CAD phenotype. Multi-way contingency tables reporting ENCODE feature and ANNOVAR annotation status with inclusion in the FDR < 5% variant list (FDR202 status) are summarized for 11 ENCODE cell types in Supplementary Table 11 and for the three CAD relevant cell types in Supplementary Table 13. Contingency table counts were modelled by a logistic multiple regression model predicting FDR202 status with independent explanatory variables HM, DHS, TFBS and genic/intergenic status. The ENCODE[83] project has previously mapped 4,492 GWAS significant SNPs from the NHGRI (June 2011) catalogue[74] to TF (12%) and DHS (34%) features in an extended dataset of 1,640 experiments. The 202 FDR variants were slightly less prevalent in these feature groups (10.4% TF and 19.8% DHS) which could reflect a CAD-specific issue or a more general consequence of our analysis being based on a subset of the ENCODE data retrieved from the Ensembl database.

Europe PMC Funders Author Manuscripts

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Majid Nikpay[#1], Anuj Goel[#2,3], Hong-Hee Won[#4,5,6,7], Leanne M Hall[#8], Christina Willenborg[#9,10], Stavroula Kanoni[#11], Danish Saleheen[#12,13], Theodosios Kyriakou[2,3], Christopher P Nelson[8,14], Jemma C Hopewell[15], Thomas R Webb[8,14], Lingyao Zeng[16,17], Abbas Dehghan[18], Maris Alver[19,20], Sebastian M Armasu[21], Kirsi Auro[22,23,24], Andrew Bjonnes[4,6], Daniel I Chasman[25,26], Shufeng Chen[27], Ian Ford[28], Nora Franceschini[29], Christian Gieger[17,30,31], Christopher Grace[2,3], Stefan Gustafsson[32,33], Jie Huang[34], Shih-Jen Hwang[35,36], Yun Kyoung Kim[37], Marcus E Kleber[38], King Wai Lau[15], Xiangfeng Lu[27], Yingchang Lu[39,40], Leo-Pekka Lyytikäinen[41,42], Evelin Mihailov[19], Alanna C Morrison[43], Natalia Pervjakova[19,22,23,24], Liming Qu[44], Lynda M Rose[25], Elias Salfati[45], Richa Saxena[4,6,46], Markus Scholz[47,48], Albert V Smith[49,50], Emmi Tikkanen[51,52], Andre Uitterlinden[18], Xueli Yang[27], Weihua Zhang[53,54], Wei Zhao[12], Mariza de Andrade[21], Paul S de Vries[18], Natalie R van Zuydam[3,55], Sonia S Anand[56], Lars Bertram[57,58], Frank Beutner[48,59], George Dedoussis[60], Philippe Frossard[13], Dominique Gauguier[61], Alison H Goodall[14,62], Omri Gottesman[39], Marc Haber[63], Bok-Ghee Han[37], Jianfeng Huang[64], Shapour Jalilzadeh[2,3], Thorsten Kessler[16,65], Inke R König[10,66], Lars Lannfelt[67], Wolfgang Lieb[68], Lars Lind[69], Cecilia M Lindgren[3,4], Marja-Liisa Lokki[70], Patrik K Magnusson[71], Nadeem H Mallick[72], Narinder Mehra[73], Thomas Meitinger[17,74,75], Fazal-ur-Rehman Memon[76], Andrew P Morris[3,77], Markku S Nieminen[78], Nancy L Pedersen[71], Annette Peters[17,30], Loukianos S Rallidis[79], Asif Rasheed[13,76], Maria Samuel[13], Svati H Shah[80], Juha Sinisalo[78], Kathleen E Stirrups[11,81], Stella Trompet[82,83], Laiyuan Wang[27,84], Khan S Zaman[85], Diego Ardissino[86,87], Eric Boerwinkle[43,88], Ingrid B Borecki[89], Erwin P Bottinger[39], Julie E Buring[25], John C Chambers[53,54,90], Rory Collins[15], L Adrienne Cupples[35,36], John Danesh[34,91], Ilja Demuth[92,93], Roberto Elosua[94], Stephen E Epstein[95], Tõnu Esko[4,19,96,97], Mary F Feitosa[89], Oscar H Franco[18], Maria Grazia Franzosi[98], Christopher B Granger[80], Dongfeng Gu[27], Vilmundur Gudnason[49,50], Alistair S Hall[99], Anders Hamsten[100], Tamara B Harris[101], Stanley L Hazen[102], Christian Hengstenberg[16,17], Albert Hofman[18], Erik Ingelsson[3,32,33,103], Carlos Iribarren[104], J Wouter Jukema[82,105,106], Pekka J Karhunen[41,107], Bong-Jo Kim[37], Jaspal S Kooner[54,90,108], Iftikhar J Kullo[109], Terho Lehtimäki[41,42], Ruth J F Loos[39,40,110], Olle Melander[111], Andres Metspalu[19,20], Winfried März[38,112,113], Colin N Palmer[55], Markus Perola[19,22,23,24], Thomas Quertermous[45,114], Daniel J Rader[115,116], Paul M Ridker[25,26], Samuli Ripatti[34,51,52], Robert Roberts[117], Veikko Salomaa[118], Dharambir K Sanghera[119,120,121], Stephen M Schwartz[122,123], Udo Seedorf[124], Alexandre F Stewart[1], David J Stott[125], Joachim Thiery[48,126], Pierre A Zalloua[63,127], Christopher J O'Donnell[35,128,129], Muredach P Reilly[116], Themistocles L Assimes[45,114], John R Thompson[130], Jeanette Erdmann[9,10], Robert Clarke[15], Hugh Watkins[2,3,133], Sekar Kathiresan[4,5,6,7,133], Ruth McPherson[1,133],

Panos Deloukas[11,131,133], Heribert Schunkert[16,17,133], Nilesh J Samani[8,14,133], and Martin Farrall[2,3,133] **for the CARDIoGRAMplusC4D Consortium**

Affiliations

[1]Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada [2]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK [3]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK [4]Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, USA [5]Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA [6]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA [7]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA [8]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK [9]Institut für Integrative und Experimentelle Genomik, Universität zu Lübeck, Lübeck, Germany [10]DZHK (German Research Center for Cardiovascular Research) partner site Hamburg-Lübeck-Kiel, Lübeck, Germany [11]William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK [12]Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA [13]Center for Non-Communicable Diseases, Karachi, Pakistan [14]NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK [15]CTSU, Nuffield Department of Population Health, University of Oxford, Oxford, UK [16]Deutsches Herzzentrum München, Technische Universität München, München, Germany [17]DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, München, Germany [18]Department of Epidemiology, Erasmus University Medical center, Rotterdam, The Netherlands [19]Estonian Genome Center, University of Tartu, Tartu, Estonia [20]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia [21]Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA [22]Department of Health, National Institute for Health and Welfare, Helsinki, Finland [23]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland [24]Diabetes & Obesity Research Program, University of Helsinki, Helsinki, Finland [25]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA [26]Harvard Medical School, Boston, Massachusetts, USA [27]State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center of Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China [28]Robertson Center for Biostatistics, University of Glasgow, Glasgow, UK [29]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, USA [30]Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany [31]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany [32]Molecular Epidemiology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden [33]Science

Europe PMC Funders Author Manuscripts

for Life Laboratory, Uppsala University, Uppsala, Sweden [34]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK [35]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA [36]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA [37]Center for Genome Science, Korea National Institute of Health, Chungcheongbuk-do, Korea [38]Vth Department of Medicine (Nephrology, Hypertensiology, Endocrinology, Diabetology, Rheumatology), Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany [39]The Charles Bronfman Institute for Personalized Medicine, The Icahn School of Medicine at Mount Sinai, New York, New York, USA [40]The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, New York, USA [41]Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland [42]Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, Finland [43]Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas, USA [44]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA [45]Department of Medicine, Division of Cardiovascular Medicine, Stanford University, Stanford, California, USA [46]Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA [47]Institute for Medical Informatics, Statistics and Epidemiology, Medical Faculty, University of Leipzig, Leipzig, Germany [48]LIFE Research Center of Civilization Diseases, Leipzig, Germany [49]Icelandic Heart Association, Kopavogur, Iceland [50]Faculty of Medicine, University of Iceland, Reykjavik, Iceland [51]Department of Public Health, University of Helsinki, Helsinki, Finland [52]Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland [53]Department of Epidemiology and Biostatistics, Imperial College London, London, UK [54]Department of Cardiology, Ealing Hospital NHS Trust, Middlesex, UK [55]Medical Research Institute, University of Dundee, Dundee, UK [56]Population Health Research Institute, Hamilton Health Sciences, Department of Medicine, McMaster University, Hamilton, Ontario, Canada [57]Platform for Genome Analytics, Institutes of Neurogenetics & Integrative and Experimental Genomics, University of Lübeck, Lübeck, Germany [58]Neuroepidemiology and Ageing Research Unit, School of Public Health, Faculty of Medicine, The Imperial College of Science, Technology, and Medicine, London, UK [59]Heart Center Leipzig, Cardiology, University of Leipzig, Leipzig, Germany [60]Department of Dietetics-Nutrition, Harokopio University, Athens, Greece [61]INSERM, UMRS1138, Centre de Recherche des Cordeliers, Paris, France [62]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK [63]Lebanese American University, School of Medicine, Beirut, Lebanon [64]Hypertension Division, Fuwai Hospital, National Center For Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China [65]Klinikum rechts der Isar, München, Germany [66]Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Lübeck, Germany [67]Department of Public Health and Caring Sciences, Geriatrics, Uppsala

University, Uppsala, Sweden [68]Institut für Epidemiologie, Christian-Albrechts Universität zu Kiel, Kiel, Germany [69]Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden [70]Transplantation Laboratory, Haartman Institute, University of Helsinki, Helsinki, Finland [71]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden [72]Punjab Institute of Cardiology, Lahore, Pakistan [73]All India Institute of Medical Sciences, New Delhi, India [74]Institut für Humangenetik, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany [75]Institute of Human Genetics, Technische Universität München, München, Germany [76]Red Crescent Institute of Cardiology, Hyderabad, Pakistan [77]Department of Biostatistics, University of Liverpool, Liverpool, UK [78]Department of Medicine, Department of Cardiology, Helsinki University Central Hospital, Helsinki, Finland [79]Second Department of Cardiology, Attikon Hospital, School of Medicine, University of Athens, Athens, Greece [80]Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA [81]Department of Haematology, University of Cambridge, Cambridge, UK [82]Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands [83]Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands [84]National Human Genome Center at Beijing, Beijing, China [85]National Institue of Cardiovascular Diseases, Karachi, Pakistan [86]Division of Cardiology, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy [87]Associazione per lo Studio della Trombosi in Cardiologia, Pavia, Italy [88]Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA [89]Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA [90]Imperial College Healthcare NHS Trust, London, UK [91]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK [92]The Berlin Aging Study II; Research Group on Geriatrics; Charité - Universitätsmedizin Berlin, Berlin, Germany [93]Institute of Medical and Human Genetics, Charité - Universitätsmedizin Berlin, Berlin, Germany [94]Grupo de Epidemiología y Genética Cardiovascular, Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain [95]MedStar Heart and Vascular Institute, MedStar Washington Hospital Center, Washington, DC, USA [96]Division of Endocrinology and Basic and Translational Obesity Research, Boston Children's Hospital, Boston, Massachusetts, USA [97]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA [98]Department of Cardiovascular Research, IRCCS Istituto di Ricerche Farmacologiche Mario Negri, Milano, Italy [99]Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK [100]Cardiovascular Genetics and Genomics Group, Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden [101]Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA [102]Cleveland Clinic, Cleveland, Ohio, USA [103]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA [104]Kaiser Permanente Division of Research, Oakland, California, USA [105]Durrer Center for Cardiogenetic

Research, Amsterdam, The Netherlands [106]Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands [107]Department of Forensic Medicine, University of Tampere School of Medicine, Tampere, Finland [108]Cardiovascular Science, National Heart and Lung Institute, Imperial College London, London, UK [109]Division of Cardiovascular Diseases, Department of Medicine, Mayo Clinic, Rochester, Minnesota, USA [110]The Mindich Child Health and Development Institute, The Icahn School of Medicine at Mount Sinai, New York, New York, USA [111]Department of Clinical Sciences, Hypertension and Cardiovascular Disease, Lund University, University Hospital Malmö, Malmö, Sweden [112]Synlab Academy, Synlab Services GmbH, Mannheim, Germany [113]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria [114]Stanford Cardiovascular Institute, Stanford University, Stanford, California, USA [115]Department of Genetics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA [116]Cardiovascular Institute, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA [117]University of Ottawa Heart Institute, Ottawa, Canada [118]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland [119]Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA [120]Department of Pharmaceutical Sciences, College of Pharmacy, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA [121]Oklahoma Center for Neuroscience, Oklahoma City, Oklahoma, USA [122]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA [123]Department of Epidemiology, University of Washington, Seattle, Washington, USA [124]Department of Prosthetic Dentistry, Center for Dental and Oral Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany [125]Institute of Cardiovascular and Medical Sciences, Faculty of Medicine, University of Glasgow, Glasgow, UK [126]Institute for Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital Leipzig, Medical Faculty, Leipzig, Germany [127]Harvard School of Public Health, Boston, Massachusetts, USA [128]National Heart, Lung and Blood Institute Division of Intramural Research, Bethesda, Maryland, USA [129]Cardiology Division, Massachusetts General Hospital, Boston, Massachusetts, USA [130]Department of Health Sciences, University of Leicester, Leicester, UK [131]Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia

Acknowledgements

We sincerely thank the participants and the medical, nursing, technical and administrative staff in each of the studies that have contributed to this project. We are grateful for support from our funders; more detailed acknowledgements are included in the Supplementary Note.

Europe PMC Funders Author Manuscripts

Author contributions

Cohort Overseeing:

D.A, E.B, I.B.B, E.P.B, J.E.B, J.C.C, R.C, L.A.C, J.D, I.D, R.E, S.E.E, T.E, M.F.F, O.H.F, M.G.F, C.B.G, D.Gu, V.G, A.S.H, A.H, T.B.H, S.L.H, C.H, A.Hofman, E.I, C.I, J.W.J, P.J.K, B-J.K, J.S.K, I.J.K, T.L, R.J.L, O.M, A.M, W.M, C.N.P, M.P, T.Q, D.J.R, P.M.R, S.R, R.R, V.S, D.K.S, S.M.S, U.S, A.F.S, D.J.S, J.T, P.A.Z, C.J.O'D, M.P.R, T.L.A, J.R.T, J.E, H.W, S.Kathiresan, R.M, P.D, H.S, N.J.S, M.F

Cohort Genotyping:

H-H.W, S.K, D.S, J.C.H, Jie Huang, M.E.K, Y.L, L-P.L, A.U, S.S.A, L.B, G.D, D.G, A.H.G, M.H, B-G.H, S.J, L.Lind, C.M.L, M-L.L, P.K.M, A.P.M, M.S.N, N.L.P, J.S, K.E.S, S.T, L.W, I.B.B, J.C.C, R.C, M.F.F, A.H, E.I, J.S.K, T.L, R.R, D.K.S, A.F.S, R.Clarke, P.D, N.J.S

Cohort Phenotyping:

D.S, J.C.H, A.D, M.A, K.A, Y.K, E.M, L.M.R, S.S.A, F.B, G.D, P.F, A.H.G, O.G, Jianfeng Huang, T.Kessler, I.R.K, L.Lannfelt, W.L, L.Lind, C.M.L, P.K.M, N.H.M, N.M, T.M, F-ur-R.M, A.P.M, N.L.P, A.P, L.S.R, A.R, M.Samuel, S.H.S, K.S.Z, D.A, J.E.B, J.C.C, R.C, R.E, C.B.G, V.G, A.S.H, A.H, S.L.H, E.I, J.W.J, P.J.K, J.S.K, I.J.K, O.M, A.M, M.P, R.R, D.K.S, A.F.S, D.J.S, P.A.Z, M.P.R, R.Clarke, S.Kathiresan, H.S, N.J.S

Cohort Data Analyst:

M.N, A.G, H-H.W, L.M.H, C.W, S.K, D.S, T.K, C.P.N, J.C.H, T.R.W, L.Z, A.D, M.A, S.M.A, K.A, A.B, D.I.C, S.C, I.F, N.F, C.G, C.Grace, S.G, Jie Huang, S-J.H, Y.K, M.E.K, K.L, X.L, Y.L, L-P.L, E.M, A.C.M, N.P, L.Q, L.M.R, E.S, R.S, M.S, A.V.S, E.T, A.U, X.Y, Weihua Zhang, Wei Zhao, M.de A, P.S.de V, N.R.van Z, M.F.F, J.R.T, M.F

Meta-analysis:

M.N, A.G, H-H.W, L.M.H, C.P.N, J.R.T, M.F

Variant annotation:

M.N, A.G, H-H.W, T.K, J.C.H, T.R.W

Paper Drafting:

M.N, A.G, H-H.W, L.M.H, T.K, J.C.H, H.W, S.Kathiresan, R.M, H.S, N.J.S, M.F

Project Steering Committee:

M.N, A.G, H-H.W, L.M.H, S.K, J.C.H, D.I.C, M.E.K, N.R.van Z, C.N.P, R.R, C.J.O'D, M.P.R, T.L.A, J.R.T, J.E, R.Clarke, H.W, S.Kathiresan, R.M, P.D, H.S, N.J.S, M.F (secretariat: J.C.H, R.Clarke)

CARDIoGRAMplusC4D Executive Committee:

J.D, D.Gu, A.H, J.S.K, R.R, H.W, S.Kathiresan, P.D, H.S, N.J.S

References

1. Kessler T, Erdmann J, Schunkert H. Genetics of coronary artery disease and myocardial infarction--2013. Curr. Cardiol. Rep. 2013; 15:368. [PubMed: 23616109]
2. O'Donnell CJ, Nabel EG. Genomics of cardiovascular disease. N. Engl. J. Med. 2011; 365:2098–109. [PubMed: 22129254]
3. CARDIoGRAMplusC4D Consortium. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat. Genet. 2013; 45:25–33. [PubMed: 23202125]
4. Coronary Artery Disease Genetics (C4D) Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat. Genet. 2011; 43:339–44. [PubMed: 21378988]
5. 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]
6. Wang F, et al. Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. Nat. Genet. 2011; 43:345–9. [PubMed: 21378986]
7. IBC 50K CAD Consortium. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. PLoS Genet. 2011; 7:e1002260. [PubMed: 21966275]
8. Clarke R, et al. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. N. Engl. J. Med. 2009; 361:2518–28. [PubMed: 20032323]
9. Bennet AM, et al. Association of apolipoprotein E genotypes with lipid levels and coronary risk. JAMA. 2007; 298:1300–11. [PubMed: 17878422]
10. Benn M, Nordestgaard BG, Grande P, Schnohr P, Tybjaerg-Hansen A. PCSK9 R46L, low-density lipoprotein cholesterol levels, and risk of ischemic heart disease: 3 independent studies and meta-analyses. J. Am. Coll. Cardiol. 2010; 55:2833–42. [PubMed: 20579540]
11. Cohen JC, Boerwinkle E, Mosley TH Jr. Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N. Engl. J. Med. 2006; 354:1264–72. [PubMed: 16554528]
12. Myocardial Infarction Genetics Consortium. A PCSK9 missense variant associated with a reduced risk of early-onset myocardial infarction. N. Engl. J. Med. 2008; 358:2299–300. [PubMed: 18499582]
13. Peloso GM, et al. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. Am. J. Hum. Genet. 2014; 94:223–32. [PubMed: 24507774]
14. Davies RW, et al. A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. Circ. Cardiovasc. Genet. 2012; 5:217–25. [PubMed: 22319020]
15. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–78. [PubMed: 17554300]
16. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol. 2010; 8:e1000294. [PubMed: 20126254]
17. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011; 88:76–82. [PubMed: 21167468]
18. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38:e164. [PubMed: 20601685]
19. Tang T, et al. hNOA1 interacts with complex I and DAP3 and regulates mitochondrial respiration and apoptosis. J. Biol. Chem. 2009; 284:5414–24. [PubMed: 19103604]
20. Chong JA, et al. REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. Cell. 1995; 80:949–57. [PubMed: 7697725]

21. Cheong A, et al. Downregulated REST transcription factor is a switch enabling critical potassium channel expression and cell proliferation. Mol. Cell. 2005; 20:45–52. [PubMed: 16209944]
22. Hao K, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. PLoS Genet. 2012; 8:e1003029. [PubMed: 23209423]
23. Salvi E, et al. Genomewide association study using a high-density single nucleotide polymorphism array and case-control design identifies a novel essential hypertension susceptibility locus in the promoter region of endothelial NO synthase. Hypertension. 2012; 59:248–55. [PubMed: 22184326]
24. Erdmann J, et al. Dysfunctional nitric oxide signalling increases risk of myocardial infarction. Nature. 2013; 504:432–6. [PubMed: 24213632]
25. Casas JP, et al. Endothelial nitric oxide synthase gene polymorphisms and cardiovascular disease: a HuGE review. Am. J. Epidemiol. 2006; 164:921–35. [PubMed: 17018701]
26. Chacon-Martinez CA, et al. The switch-associated protein 70 (SWAP-70) bundles actin filaments and contributes to the regulation of F-actin dynamics. J. Biol. Chem. 2013; 288:28687–703. [PubMed: 23921380]
27. Zeller T, et al. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. PLoS One. 2010; 5:e10693. [PubMed: 20502693]
28. Fairfax BP, et al. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. Science. 2014; 343:1246949. [PubMed: 24604202]
29. Grundberg E, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat. Genet. 2012; 44:1084–9. [PubMed: 22941192]
30. Ashcroft GS, et al. Mice lacking Smad3 show accelerated wound healing and an impaired local inflammatory response. Nat. Cell Biol. 1999; 1:260–6. [PubMed: 10559937]
31. Samani NJ, et al. Genomewide association analysis of coronary artery disease. N. Engl. J. Med. 2007; 357:443–53. [PubMed: 17634449]
32. Silvestre JS, et al. Lactadherin promotes VEGF-dependent neovascularization. Nat. Med. 2005; 11:499–506. [PubMed: 15834428]
33. Hanayama R, et al. Identification of a factor that links apoptotic cells to phagocytes. Nature. 2002; 417:182–7. [PubMed: 12000961]
34. Miyata K, et al. Elevated mature macrophage expression of human ABHD2 gene in vulnerable plaque. Biochem. Biophys. Res. Commun. 2008; 365:207–13. [PubMed: 17980156]
35. Jain M, Bhat GP, Vijayraghavan K, Inamdar MS. Rudhira/BCAS3 is a cytoskeletal protein that controls Cdc42 activation and directional cell migration during angiogenesis. Exp. Cell Res. 2012; 318:753–67. [PubMed: 22300583]
36. Kim JY, Ahn HJ, Ryu JH, Suk K, Park JH. BH3-only protein Noxa is a mediator of hypoxic cell death induced by hypoxia-inducible factor 1alpha. J. Exp. Med. 2004; 199:113–24. [PubMed: 14699081]
37. Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. Nat. Genet. 2013; 45:1274–83. [PubMed: 24097068]
38. Lango Allen H, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010; 467:832–8. [PubMed: 20881960]
39. Morris AP, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat. Genet. 2012; 44:981–90. [PubMed: 22885922]
40. Scott RA, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat. Genet. 2012; 44:991–1005. [PubMed: 22885924]
41. Speliotes EK, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat. Genet. 2010; 42:937–48. [PubMed: 20935630]
42. Pearce LR, et al. KSR2 mutations are associated with obesity, insulin resistance, and impaired cellular fuel oxidation. Cell. 2013; 155:765–77. [PubMed: 24209692]
43. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. Curr. Opin. Genet. Dev. 2009; 19:212–9. [PubMed: 19481926]

44. Lettre G, Lange C, Hirschhorn JN. Genetic model testing and statistical power in population-based association studies of quantitative traits. Genet. Epidemiol. 2007; 31:358–62. [PubMed: 17352422]
45. Do R, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature. 2015; 518:102–6. [PubMed: 25487149]
46. TG and HDL Working Group of the Exome Sequencing Project. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. N. Engl. J. Med. 2014; 371:22–31. [PubMed: 24941081]
47. Myocardial Infarction Genetics Consortium Investigators. Inactivating mutations in NPC1L1 and protection from coronary heart disease. N. Engl. J. Med. 2014; 371:2072–82. [PubMed: 25390462]
48. Maurano MT, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012; 337:1190–5. [PubMed: 22955828]
49. Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. Nature. 2011; 473:317–25. [PubMed: 21593864]
50. Reilly MP, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. Lancet. 2011; 377:383–92. [PubMed: 21239051]
51. Dichgans M, et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. Stroke. 2014; 45:24–36. [PubMed: 24262325]
52. Keating BJ, et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. PLoS One. 2008; 3:e3583. [PubMed: 18974833]
53. Voight BF, et al. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet. 2012; 8:e1002793. [PubMed: 22876189]
54. Schunkert H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat. Genet. 2011; 43:333–8. [PubMed: 21378990]
55. Miyata K, et al. Increase of smooth muscle cell migration and of intimal hyperplasia in mice lacking the alpha/beta hydrolase domain containing 2 gene. Biochem. Biophys. Res. Commun. 2005; 329:296–304. [PubMed: 15721306]
56. Bobik A. Transforming growth factor-betas and vascular disorders. Arterioscler. Thromb. Vasc. Biol. 2006; 26:1712–20. [PubMed: 16675726]
57. Mallat Z, et al. Inhibition of transforming growth factor-beta signaling accelerates atherosclerosis and induces an unstable plaque phenotype in mice. Circ. Res. 2001; 89:930–4. [PubMed: 11701621]
58. Yang Z, et al. Infarct-sparing effect of A2A-adenosine receptor activation is due primarily to its action on lymphocytes. Circulation. 2005; 111:2190–7. [PubMed: 15851591]
59. Aziz M, Jacob A, Matsuda A, Wang P. Review: milk fat globule-EGF factor 8 expression, function and plausible signal transduction in resolving inflammation. Apoptosis. 2011; 16:1077–86. [PubMed: 21901532]

Method References

60. Yang J, et al. Genomic inflation factors under polygenic inheritance. Eur. J. Hum. Genet. 2011; 19:807–12. [PubMed: 21407268]
61. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat. Genet. 2012; 44:955–9. [PubMed: 22820512]
62. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]
63. Magi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. BMC bioinformatics. 2010; 11:288. [PubMed: 20509871]
64. Cochran WG. The Combination of Estimates from Different Experiments. Biometrics. 1954; 10:101–129.

65. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat. Med. 2002; 21:1539–58. [PubMed: 12111919]
66. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control. Clin. Trials. 1986; 7:177–88. [PubMed: 3802833]
67. Newson RB. Frequentist q-values for multiple-test procedures. The Stata Journal. 2010; 10:568–84.
68. Benjamini Y, Yekutieli D. The control of the false-discovery rate in multiple testing under dependency. Ann. Stats. 2001; 29:1165–88.
69. Yang J, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat. Genet. 2012; 44:369–75. S1–3. [PubMed: 22426310]
70. So HC, Gui AH, Cherny SS, Sham PC. Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. Genet. Epidemiol. 2011; 35:310–7. [PubMed: 21374718]
71. Heid IM, et al. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat. Genet. 2010; 42:949–60. [PubMed: 20935629]
72. International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011; 478:103–9. [PubMed: 21909115]
73. Wain LV, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat. Genet. 2011; 43:1005–11. [PubMed: 21909110]
74. Welter D, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014; 42:D1001–6. [PubMed: 24316577]
75. Fehrmann RS, et al. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. PLoS Genet. 2011; 7:e1002197. [PubMed: 21829388]
76. Garnier S, et al. Genome-wide haplotype analysis of cis expression quantitative trait loci in monocytes. PLoS Genet. 2013; 9:e1003240. [PubMed: 23382694]
77. Gibbs JR, et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. PLoS Genet. 2010; 6:e1000952. [PubMed: 20485568]
78. Liang L, et al. A cross-platform analysis of 14,177 expression quantitative trait loci derived from lymphoblastoid cell lines. Genome Res. 2013; 23:716–26. [PubMed: 23345460]
79. Westra HJ, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. 2013; 45:1238–43. [PubMed: 24013639]
80. Busch SJ, Barnhart RL, Martin GA, Flanagan MA, Jackson RL. Differential regulation of hepatic triglyceride lipase and 3-hydroxy-3-methylglutaryl-CoA reductase gene expression in a human hepatoma cell line, HepG2. J. Biol. Chem. 1990; 265:22474–9. [PubMed: 2176219]
81. Park HJ, et al. Human umbilical vein endothelial cells and human dermal microvascular endothelial cells offer new insights into the relationship between lipid metabolism and angiogenesis. Stem Cell Rev. 2006; 2:93–102. [PubMed: 17237547]
82. Durrani S, Konoplyannikov M, Ashraf M, Haider KH. Skeletal myoblasts for cardiac repair. Regen. Med. 2010; 5:919–32. [PubMed: 21082891]
83. Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]

Europe PMC Funders Author Manuscripts

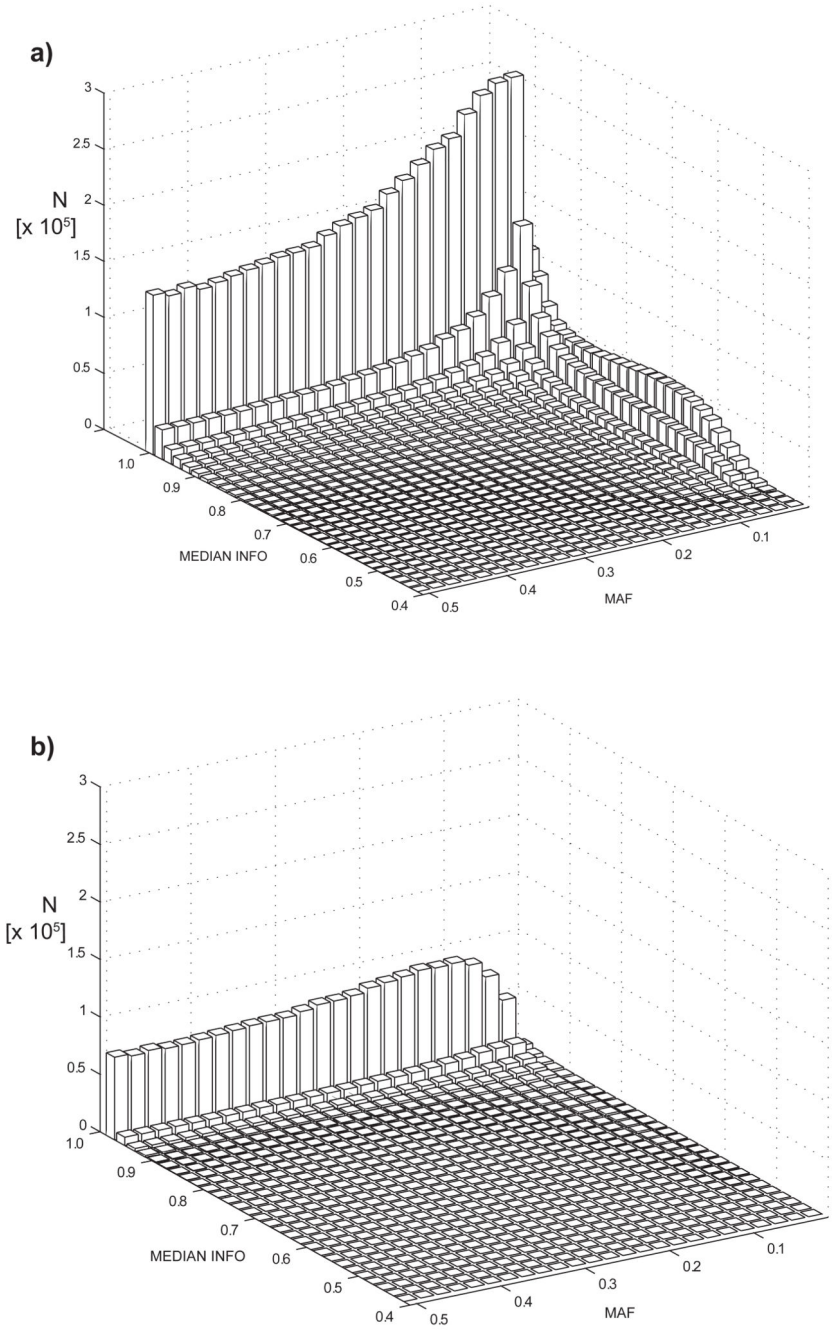

Figure 1.

Spectrum of minor allele frequencies (MAF) and median imputation quality (MEDIAN INFO) showing the number (N) of variants in each bin (**a**) shows the distribution for the 9.4M 1000 Genomes phase1v3 variants (**b**) shows the distribution for 2.5M HapMap2 SNPs. Imputation quality was calculated as the median of the respective values in up to 48 contributing studies; the imputation quality for genotyped variants was set equal to 1.0. The 1000 Genomes training set included more low frequency variants, many of which have imputation qualities > 0.9.

Europe PMC Funders Author Manuscripts

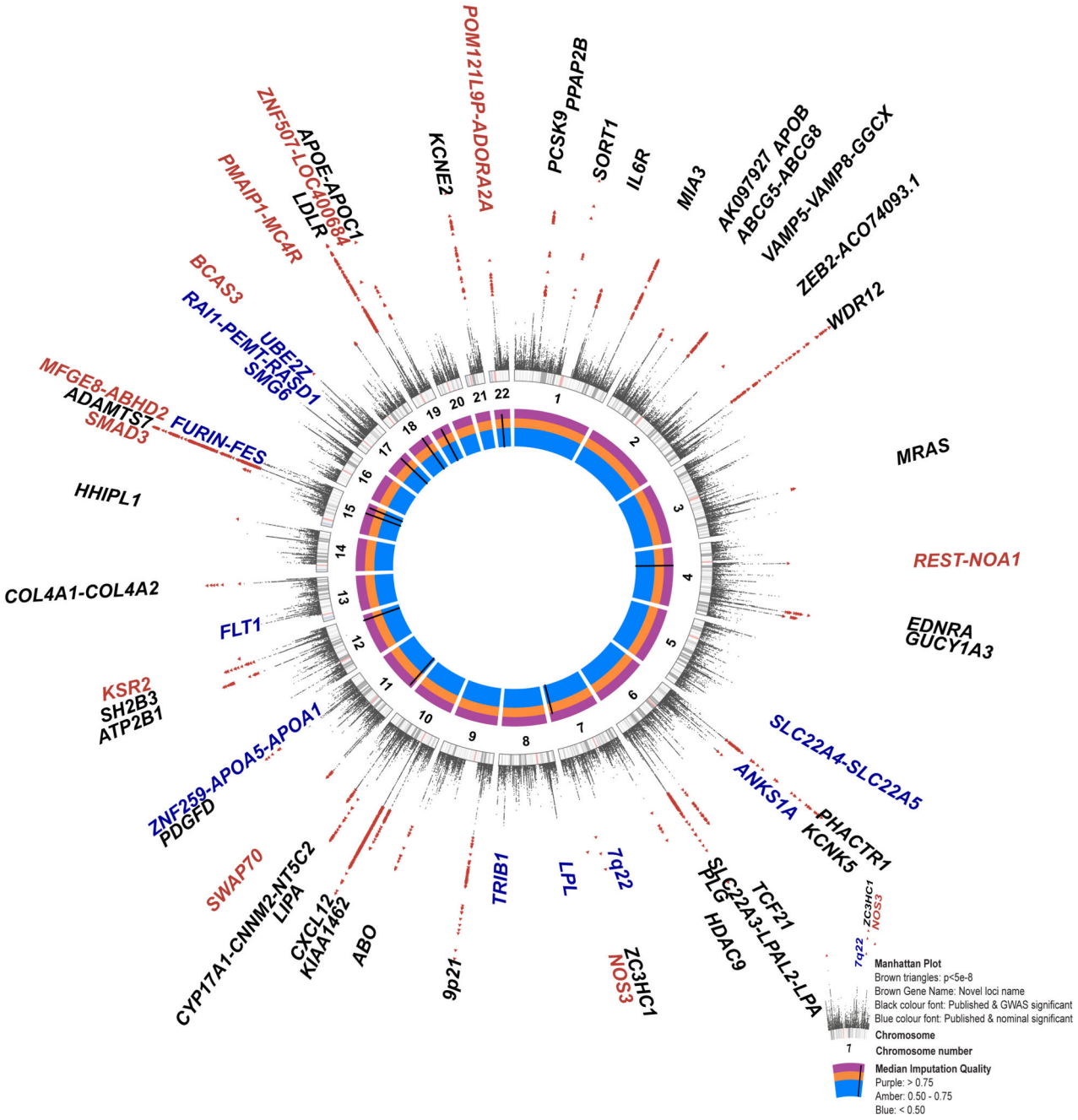

Figure 2.

A circular Manhattan plot summarizing the 1000 Genomes CAD association results. The meta-analysis statistics have been adjusted for over-dispersion (before double genomic control, lambda = 1.18); over-dispersion is predicted to be a regular feature in GWAS under the polygenic inheritance model[60]. The association statistics have been capped to $P = 1 \times 10^{-20}$. Genome-wide significant variants ($P < 5 \times 10^{-8}$) are indicated by red triangles. Novel CAD loci are named in red (Table 1). Previously reported loci showing genome-wide significance are shown in black and those showing nominal significance ($P < 0.05$) in our

meta-analysis in blue (Supplementary Table 2). The inner track (see inset) shows the imputation quality score of the lead variants of the novel loci. The middle track shows numbered chromosome ideograms with the centromeres indicated by pink bars.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

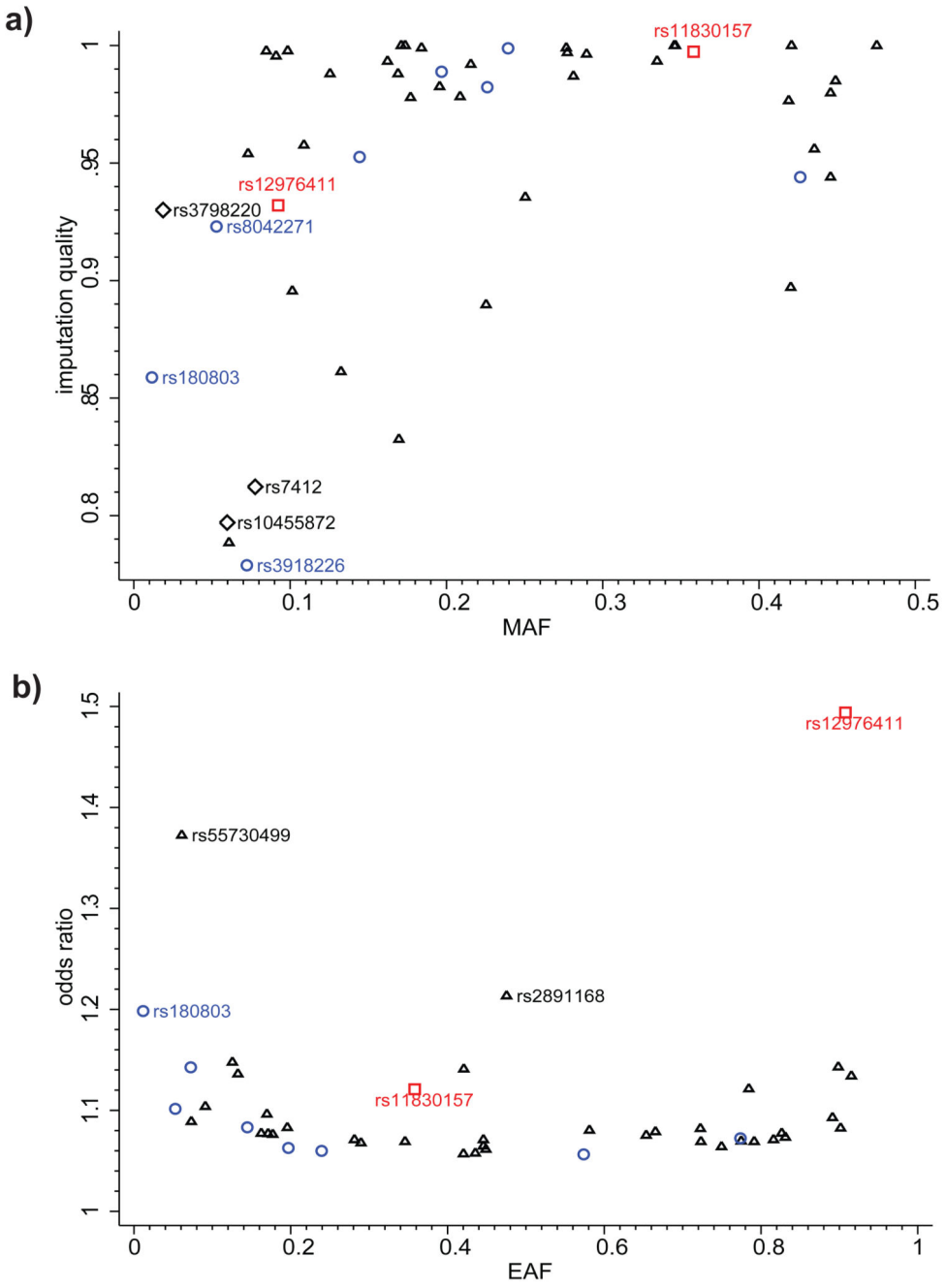

Figure 3.

Imputation quality and effect size of lead variants at 46 genome-wide significant loci. **(a)** Imputation quality and minor allele frequency (MAF) for lead variants at 46 genome-wide significant susceptibility loci. Blue circles indicate novel additive loci, red squares - novel recessive loci, black triangles - previously mapped additive loci, black diamonds - key SNPs in *LPA* and *APOE*. Imputation quality and MAF were calculated as the median of the respective values in up to 48 contributing studies; the imputation quality for studies with genotype data was fixed at 1.0. **(b)** Odds ratios and effect allele frequency (EAF) for lead

variants at 46 genome-wide significant loci. Blue circles indicate novel additive loci; red squares - novel recessive loci, black triangles - previously mapped additive loci. SNPs rs55730499 and rs2891168 are lead variants in the *LPA* and chromosome 9p21 susceptibility loci. EAF was calculated as the median value in up to 48 contributing studies.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

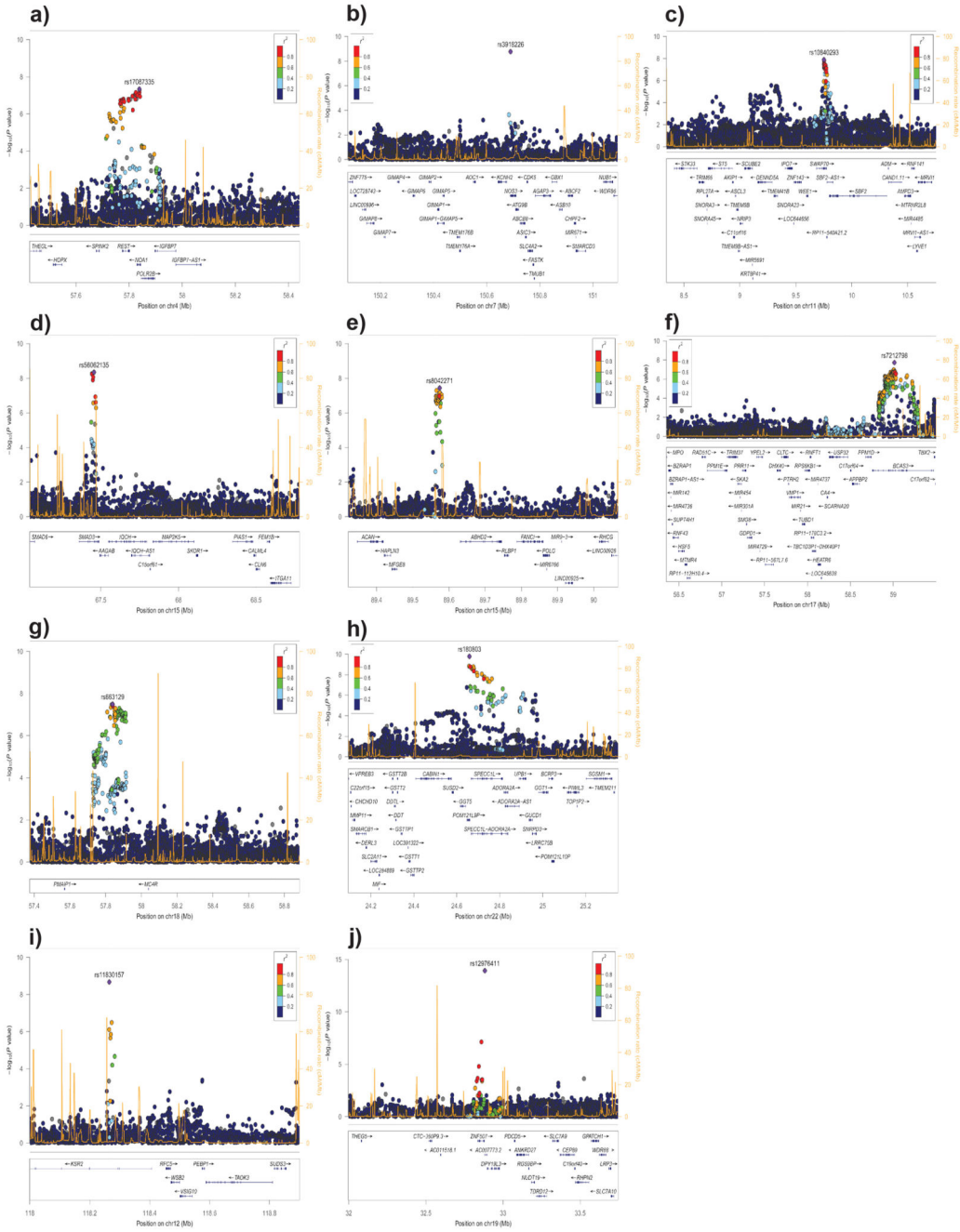

Figure 4.

Regional association plots of the eight additive (**a–h**) and two recessive (**i–j**) novel CAD loci. The association statistics have been adjusted for over-dispersion following meta-analysis (genomic control parameter 1.18 for the additive and 1.05 for the recessive models). Linkage disequilibrium (r^2) calculations were based on the combined 1000 Genomes phase 1 v3 training dataset. Genomic coordinates in mega-base pairs (Mb) refer to the hg19 sequence assembly.

Europe PMC Funders Author Manuscripts

Table 1

Ten new CAD loci

									Association model			
									Additive		Recessive	
Lead variant	**Locus name**	**Chr.**	**A1/A2**	**Effect allele (A1) freq.**	**Imputation quality**	**I^2**	**Heterogeneity *P***	***n* studies**	**OR (95% CI)**	***P***	**OR (95% CI)**	***P***
rs17087335	*REST-NOA1*	4	T/G	0.21	0.99	0.20	0.11	48	**1.06 (1.04-1.09)**	**4.60E-08**	1.11 (1.05-1.17)	3.30E-04
rs3918226	*NOS3*	7	T/C	0.06	0.78	0.15	0.19	45	**1.14 (1.09-1.19)**	**1.70E-09**	1.26 (0.99-1.60)	5.96E-02
rs10840293	*SWAP70*	11	A/G	0.55	0.94	0.17	0.16	47	**1.06 (1.04-1.08)**	**1.30E-08**	1.05 (1.02-1.09)	1.51E-03
rs56062135	*SMAD3*	15	C/T	0.79	0.98	0.00	0.67	46	**1.07 (1.05-1.10)**	**4.50E-09**	1.17 (1.10-1.25)	8.88E-07
rs8042271	*MFGE8-ABHD2*	15	G/A	0.9	0.93	0.16	0.19	46	**1.10 (1.06-1.14)**	**3.70E-08**	1.25 (1.13-1.37)	7.27E-06
rs7212798	*BCAS3*	17	C/T	0.15	0.95	0.14	0.21	48	**1.08 (1.05-1.11)**	**1.90E-08**	1.17 (1.07-1.28)	6.12E-04
rs663129	*PMAIP1-MC4R*	18	A/G	0.26	1.00	0.00	0.6	47	**1.06 (1.04-1.08)**	**3.20E-08**	1.11 (1.06-1.17)	7.15E-06
rs180803	*POM121L9P-ADORA2A*	22	G/T	0.97	0.86	0.00	0.67	41	**1.20 (1.13-1.27)**	**1.60E-10**	N/A	N/A
rs11830157	*KSR2*	12	G/T	0.36	0.99	0.14	0.22	42	1.04 (1.02-1.06)	3.90E-04	**1.12 (1.08-1.16)**	**2.12E-09**
rs12976411	*ZNF507-LOC400684*	19	T/A	0.09	0.93	0.50	5.09E-04	34	0.95 (0.92-0.99)	9.10E-03	**0.67 (0.60-0.74)**	**1.18E-14**

Association results are presented for two inheritance models; results in bold text indicate the discovery association model. *P* values have been adjusted for over-dispersion following meta-analysis.

Heterogeneity *P* values are for the respective discovery association model. Chr., chromosome; A1, effect allele; A2, non-effect allele; freq., frequency; I^2, heterogeneity inconsistency index; OR, odds ratio; CI, confidence interval; N/A, not available due to insufficient numbers (<60%) of studies for reliable results. The *n* studies column shows the number of studies participated in the discovery result where up to 48 studies participated in the additive model and up to 43 studies in the recessive model meta-analysis.